J Intell Robot Syst (2011) 61:157–168
DOI 10.1007/s10846-010-9489-5

A Visual Global Positioning System for Unmanned Aerial Vehicles Used in Photogrammetric Applications

Andrea Cesetti · Emanuele Frontoni · Adriano Mancini · Andrea Ascani · Primo Zingaretti · Sauro Longhi

Received: 1 February 2010 / Accepted: 1 September 2010 / Published online: 23 November 2010

Abstract The combination of photogrammetric aerial and terrestrial recording methods can provide new opportunities for photogrammetric applications. A UAV (Unmanned Aerial Vehicle), in our case a helicopter system, can cover both the aerial and quasi-terrestrial image acquisition methods. A UAV can be equipped with an on-board high resolution camera and a priori knowledge of the operating area where to perform photogrammetric tasks. In this general scenario our paper proposes vision-based techniques for localizing a UAV. Only natural landmarks provided by a feature tracking algorithm will be considered, without the help of visual beacons or landmarks with known positions. The novel idea is to perform global localization, position tracking and localization failure recovery (kidnapping) based only on visual matching between current view and available georeferenced satellite images. The matching is based on SIFT features and the system estimates the position of the UAV and its altitude on the base of the reference image. The vision system replaces the GPS signal combining position information from visual odometry and georeferenced imagery. Georeferenced satellite or aerial images must be available on-board beforehand or downloaded during the flight. The growing availability

A. Cesetti (✉) · E. Frontoni · A. Mancini · A. Ascani · P. Zingaretti · S. Longhi
Dipartimento di Ingegneria Informatica, Gestionale e dell'Automazione,
Università Politecnica delle Marche, Ancona, Italy
e-mail: cesetti@diiga.univpm.it

E. Frontoni
e-mail: frontoni@diiga.univpm.it

A. Mancini
e-mail: mancini@diiga.univpm.it

A. Ascani
e-mail: ascani@diiga.univpm.it

P. Zingaretti
e-mail: zinga@diiga.univpm.it

S. Longhi
e-mail: sauro.longhi@univpm.it

of high resolution satellite images (e.g., provided by Google Earth or other local information sources) makes this topic very interesting and timely. Experiments with both synthetic (i.e., taken from satellites or datasets and pre elaborated) and real world images have been performed to test the accuracy and the robustness of our method. Results show sufficient performance if compared with common GPS systems and give a good performance also in the altitude estimation, even if in this last case there are only preliminary results.

Keywords Visual global positioning system · Unmanned aerial vehicles · Photogrammetric applications · Natural landmark visual matching

1 Introduction

Over the last few years civilian remote sensing applications have greatly increased owing to progress in satellite and/or aerial image acquisition quality (resolution) and accuracy. They range from photogrammetric recording of archeological sites [10], rapid emergency response operations [14] and environmental monitoring to inspection of rooftops for breakings and 3D vector map production [15].

Image acquisition by satellites and manned aircrafts shows several limitations, like high launch/flight costs, slow and weather-dependent data collection, restricted maneuverability and limited availability, and so UAV image acquisition technologies have been developed in recent years. In particular, more and more applications of UAV systems have become common for photogrammetric applications, where acquired images need to be georeferenced to be combined with existing data in Geographical Information Systems (GIS). This development can be mainly explained by: 1) a cheap and easy engineering organization of aerial photogrammetric surveys; 2) the opportunity of performing very low altitude aerial photography at cloudy days; 3) getting many images of objects (e.g., city buildings) from different directions by complicated flights; 4) the spreading of low cost combined GPS/INS systems, which are necessary to navigate the UAV with high precision to the predicted acquisition points.

The main reason that prevents a still larger use of UAV systems is the safety issue. State of the art UAV systems are still not able to guarantee an acceptable level of safety to convince aviation authorities to authorize the use of such a system in populated areas (except in rare cases such as war zones). There are several problems that have to be solved before unmanned aircrafts can be introduced in the civilian airspace. One of them is the GPS integrity problem [7].

Autonomous UAVs usually rely on a GPS position signal that, combined with Inertial Measurement Unit (IMU) data, provides high-rate and drift-free state estimation, suitable for control purposes. Small UAVs are usually equipped with low performance IMUs due to their limited payload capabilities. In such platforms, the loss of the GPS signal, even for a few seconds, can be catastrophic due to the high drift rate of the IMU installed on-board. Besides, the GPS signal can become unreliable when operating close to obstacles, due to multi-path reflections, and is quite vulnerable to jamming (especially for a GPS operating on civilian frequencies). In addition, GPS remote sensing measurement accuracy [13] could result a problem, being affected by a number of factors, including satellite positions, noise in the radio signal, atmospheric conditions and natural barriers to the signal. Noise caused by interference from something near the receiver can create an error between 1 to 10 m. Objects such mountains or buildings between the satellite and the receiver can also produce accuracy errors, sometimes up to 30 m.

Because, for all these reasons, UAVs that rely blindly on a GPS signal are quite vulnerable, this paper proposes a navigation system that can cope with short and long term GPS outages. In particular, a novel Visual Global Positioning System (VGPS) for Unmanned Aerial Vehicles, which is useful when the GPS signal is lost or in the case of malfunctioning of on-board navigation instrumentation, is proposed. It uses only a vision system based on a passive monocular video camera. The vision system replaces the GPS signal combining position information from visual odometry and georeferenced imagery. We are particularly interested in low altitude UAV photogrammetric applications, like the evaluation of the impact of calamities (e.g., damage assessment after an inundation or an earthquake), map update and environment monitoring. Georeferenced satellite or aerial images must be available on-board of the UAV beforehand or downloaded during the flight. The growing availability of high resolution satellite images (e.g., provided by Google Earth or other local information sources) makes this topic very interesting and timely.

Our method assumes that the imaged scene is taken by a downward pointing camera while the UAV is flying in a quasi-hover condition. Under this hypothesis it is reasonable to compare the taken image with a part from the known database (e.g., satellite or high altitude images). Moreover the UAV is required to fly at relative high altitude compared to the objects on the ground. This assumption is not restrictive because small UAV inspection and photogrammetric applications are supposed to be performed at a high enough altitude.

The estimated position does not present drifts along the sequence of images. The visual system is based on point feature matching and a novel and robust filtering method based on Feature Group Matching (FGM) [2]. Mosaicing helps to reduce the registration errors and, hence, to increase speed performances of the position estimation.

The paper is organized as follows. Section 2 presents related works on UAV position estimation based on vision. Section 3 presents the method implemented to dynamically localize the UAV while improving the accuracy and robustness of the estimation. Section 4 presents the experimental data and the results obtained. Finally, the conclusions and future trends mentioned in Section 5 complete the paper.

2 Related Works

Research on vision applied to UAV position estimation starts in the nineties; in [1] it is described a vision-based odometer that allowed deriving the relative helicopter position and velocity in real time by means of stereo vision. The paper also demonstrated autonomous tracking capabilities of moving objects by using only on-board processing power.

Also vision-based autonomous landing of UAV systems has been actively researched [18]. In [9] Dickmanns and Schell presented some results of the possible use of vision for landing an airplane. Systems based on artificial landmarks and structured light are presented in [25, 26]. The BEAR project is a good example of vision systems for autonomous landing of UAVs that use vision-based pose estimation relative to a planar landing target and vision-based landing of an aerial vehicle on a moving deck [21, 23]. A method based on multiple view geometry is used to compute the real motion of one UAV with respect to a planar landing target. An artificial target allows to establish quick matches and to solve the scale problem. In [12] a strategy and an algorithm relying on image processing to search the ground for a safe landing spot is presented. Vision-based techniques for landing on an artificial helipad of known shape are also presented in [19, 20], where the case of landing on a slow moving helipad is considered. In [22] the landing strategies of bees are used to devise a vision system based on optical flow for UAVs.

Corke et. al [8] have analyzed the use of stereo vision for height estimation in small size helicopters. In [3] a visual odometer for aerial vehicles using monocular image sequences is presented, but no error estimation is provided by the algorithm, and the approach is limited to planar scenes. In [4] it is shown how a mosaic can be used in aerial vehicles to partially correct the drift associated to odometric approaches. In [24] the authors presented an Extended Kalman Filter approach that combines GPS measurements with image features obtained from a known artificial target for position estimation.

Recently, in [6] the authors proposed a visual system based on feature extraction working on safe landing and on general control strategy for UAVs [5, 17]. The general idea was the use of a multi-purpose feature based approach to cope with navigation, landing and cooperative UAV tasks.

This paper extends our previous works by proposing novel applications in photogrammetric scenarios: a visual GPS based on previously known geo-referenced images, which are usually available in this field.

3 Overview of the Method

The proposed method is based on image matching between a previously known database of georeferenced images and the current UAV view. The matching is performed using the SIFT point feature extractor [16] and a couple of filtering techniques: the first is based on the classic RANSAC [11] approach, the second is the novel FGM [2] method developed by our group.

The x and y positions of the UAV are computed on the basis of image matching results with respect to the georeferenced images. An estimation of the altitude, based on simple triangulation, given the altitude of the georeferenced images and camera parameters, is also provided.

Here following details about the used methods for image matching and filtering of outliers are given.

3.1 SIFT Detector

SIFT [16] consists of four major stages: scale-space extrema detection, keypoint localization, orientation assignment and keypoint descriptor. In the first stage, to improve the computation speed, on the contrary of the Gaussian kernel the Difference-of-Gaussian kernel is used to identify scale and orientation invariant potential interest points. In the keypoint localization step, the low contrast points are rejected and edge responses eliminated; the Hessian matrix is used to compute the principal curvatures and to eliminate the keypoints that have a ratio between the principal curvatures greater than ratio. An orientation histogram is formed from the gradient orientations of sample points within a region around the keypoint to get orientation assignment. According to experiments we performed, best results are achieved with a 4*4 array of histograms with 8 orientation bins. So the resulting dimension of the adopted SIFT descriptor is 4*4*8 = 128.

3.2 RANSAC

A first common evaluation and filtering method is RANSAC [11], which is used to reject inconsistent matches between features. The goal is to obtain the inliers and reject outliers at the same time. The probability that the algorithm never selects a set of m points that all are inliers is (1 – p):

$$1 - p = \left(1 - w^m\right)^k$$

where m is the least number of points needed for estimating a model, k is the number of samples required, w is the probability that the RANSAC algorithm selects inliers from the input data set. The RANSAC repeatedly estimates a set of models drawn randomly from the input set. At every step RANSAC tries to optimize the correspondence to find a dominant model that is then used to filter out incorrect matchings.

3.3 Feature Group Matching

In this approach we propose the use of feature spatial relations to solve the outlier filtering problem. In practice, we create a feature group for each feature in the matched images I_1 and I_2 and represent it with a vector, the group descriptor. Then we perform easy and fast comparisons between group descriptors, one from I1 and the other from I_2, to determine stable features. This approach lightly increases the computational time, but leads to a better overall image matching by filtering out most incorrect matchings. In this way FMG [2] is used as an advanced filter to validate each feature matching.

Referring to Fig. 1, a group is a set of n features, with $n > 1$. The location of two of these features is essential for the definition of the group. The location C of the feature on which the group is built is considered the centre of the group. The location P of the nearest feature to C permits to split the space around C to describe feature spatial relations. If d is the distance between C and P, M concentric circumferences and L radial half-lines are defined. Each circumference is described by the same centre C and radius given by $r_i = d / i$, where i is the index of the circumference. A descriptor for the group, which tries to represent different spatial displacements of point features around the group centre, is defined in [2].

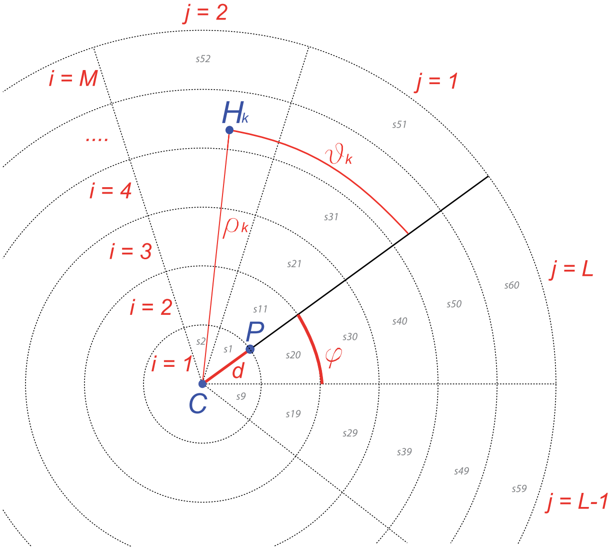

Fig. 1 Feature group description

4 Tests and Results

This section shows the tests performed and the results related to localization obtained with the techniques described above. Tests were performed offline using real data consisting of two series of images that cover the same area (more than 200 km^2 in width).

The first set of data, used as reference, consists of georeferenced satellite images with a spatial resolution of 4 m per pixel. The second set of data consists of orthorectified aerial images with a spatial resolution of 1 m per pixel. These high-resolution images have been appropriately cropped and rotated to simulate the visual information available for the aircraft during the flight.

The two sets of images were recorded in different years and different seasons and therefore differ considerably with regard to colours and shapes of the geographical features (e.g., boundaries of cultivated lands, types of cultivations, new buildings). Figure 2 shows some of the typical differences outlined above.

The whole flight strip, representing the a priori knowledge of the environment, has been divided into smaller images (Fig. 3) to keep limited the computational load for the on-board system. Tests have been performed using sub-images of 256 × 256 or 512 × 512 pixels. For aerial images the size of 512 × 512 pixels has been usually chosen. Larger images (1,024 × 1,024 pixels) have been resized before being processed by the feature extractor.

Several tests have shown the effectiveness of the approach; the presence of invariant features in the image, like buildings, particularly increases the number of features matched (Fig. 4). When the characteristics of the ground are not so favourable, the problem due to the greater presence in percentage of erroneous

Fig. 2 Comparison of an aerial image and a satellite image of the same area. Significant differences in colours and boundaries of pieces of ground can be easily noted

matchings is overcome by a filtering. In this case the algorithm based on group-matching guarantees a better performance in terms of time and effectiveness (Fig. 5).

The following tests address the "kidnapping" problem: the system is required to locate the high definition aerial image by comparing it with a group of sub-images

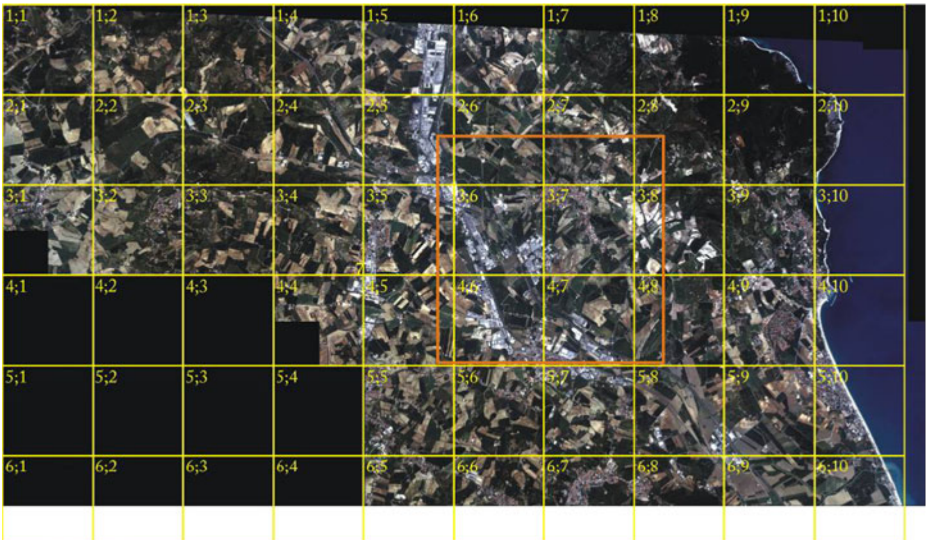

Fig. 3 The whole satellite image divided into a matrix of sub-images. The *orange square* identifies the area considered for the tests

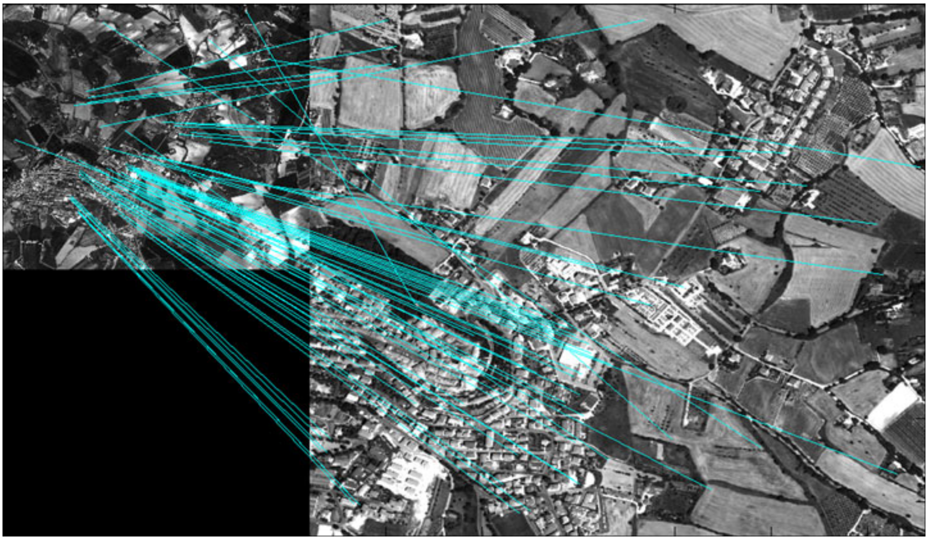

Fig. 4 The matching between the satellite image (*on the left*) and the aerial one

relating to an area in the neighbourhood of the flight area. Figure 6 shows the result of a test performed using SIFT features and RANSAC filtering. The known map is constituted by a matrix of 10 × 20 sub-images of 512 × 512 pixels. The comparison was performed among an aerial image (256 × 256 pixels) and 25 images of the map. The pink square indicates the sub-image that presents the higher number of matchings that were considered correct by the system.

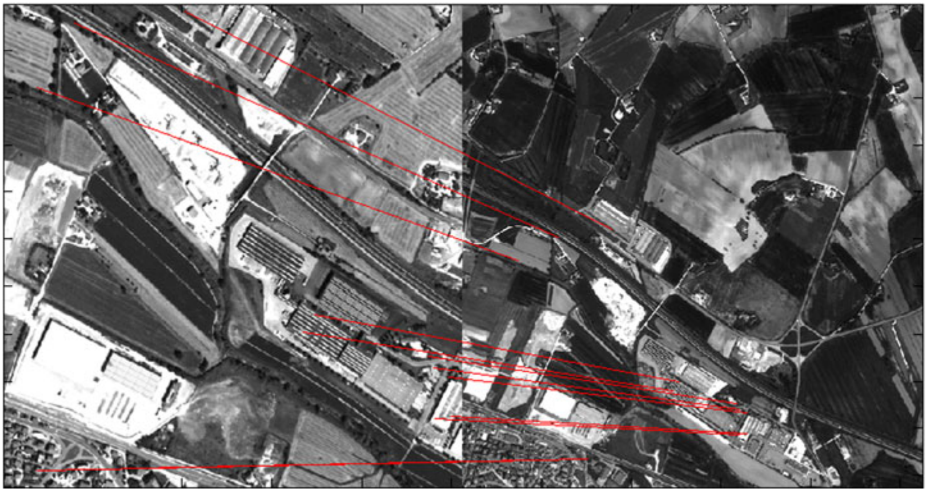

Fig. 5 The matching results after FGM filtering

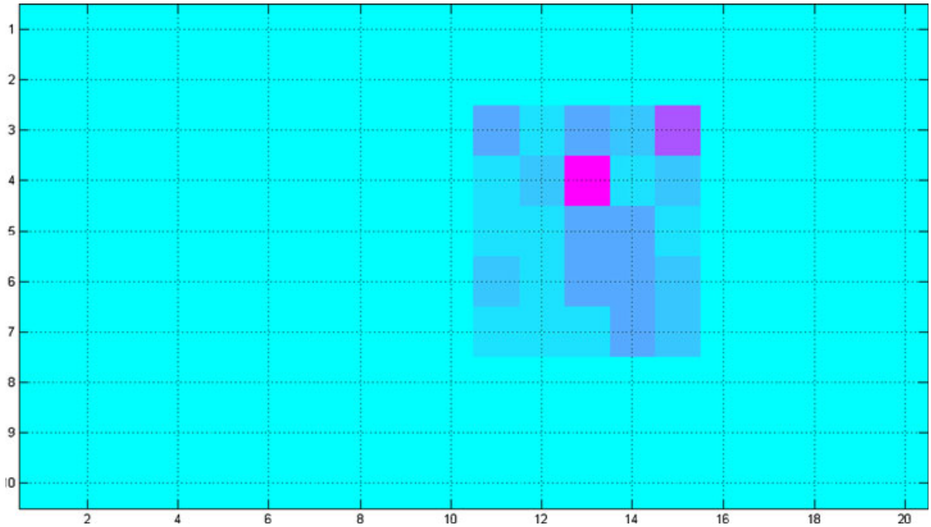

Fig. 6 The solution of a kidnapping problem: the *pink square* represents the map image that matches with the aerial one

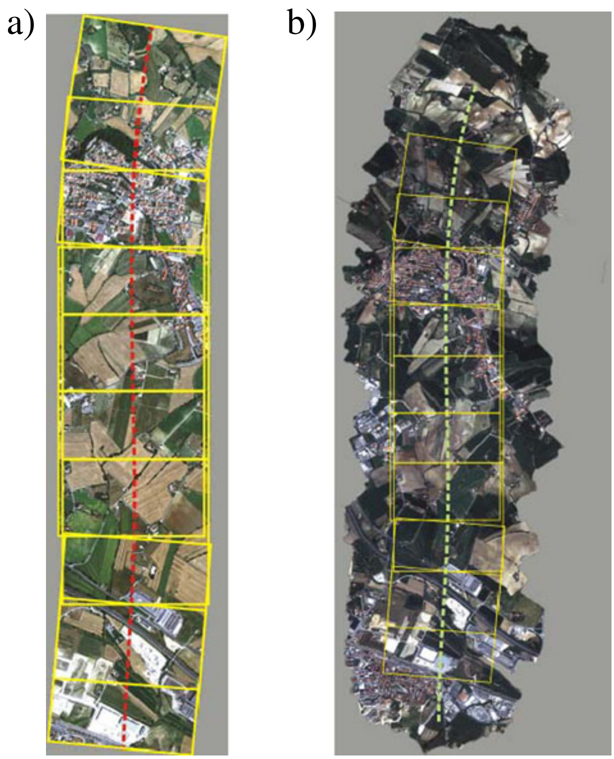

Fig. 7 The simulated path composed by the aerial images (**a**) and reported on the satellite geo-referenced image (**b**)

Further tests simulate the visual-GPS operating during a flight. Several images were collected, composing a possible path followed by the aircraft (Fig. 7). Each of them was processed by the system and located respect to the known map. Exploiting the geo-references associated to the pixels of the satellite images a 2D absolute localization of the aircraft was easily computed.

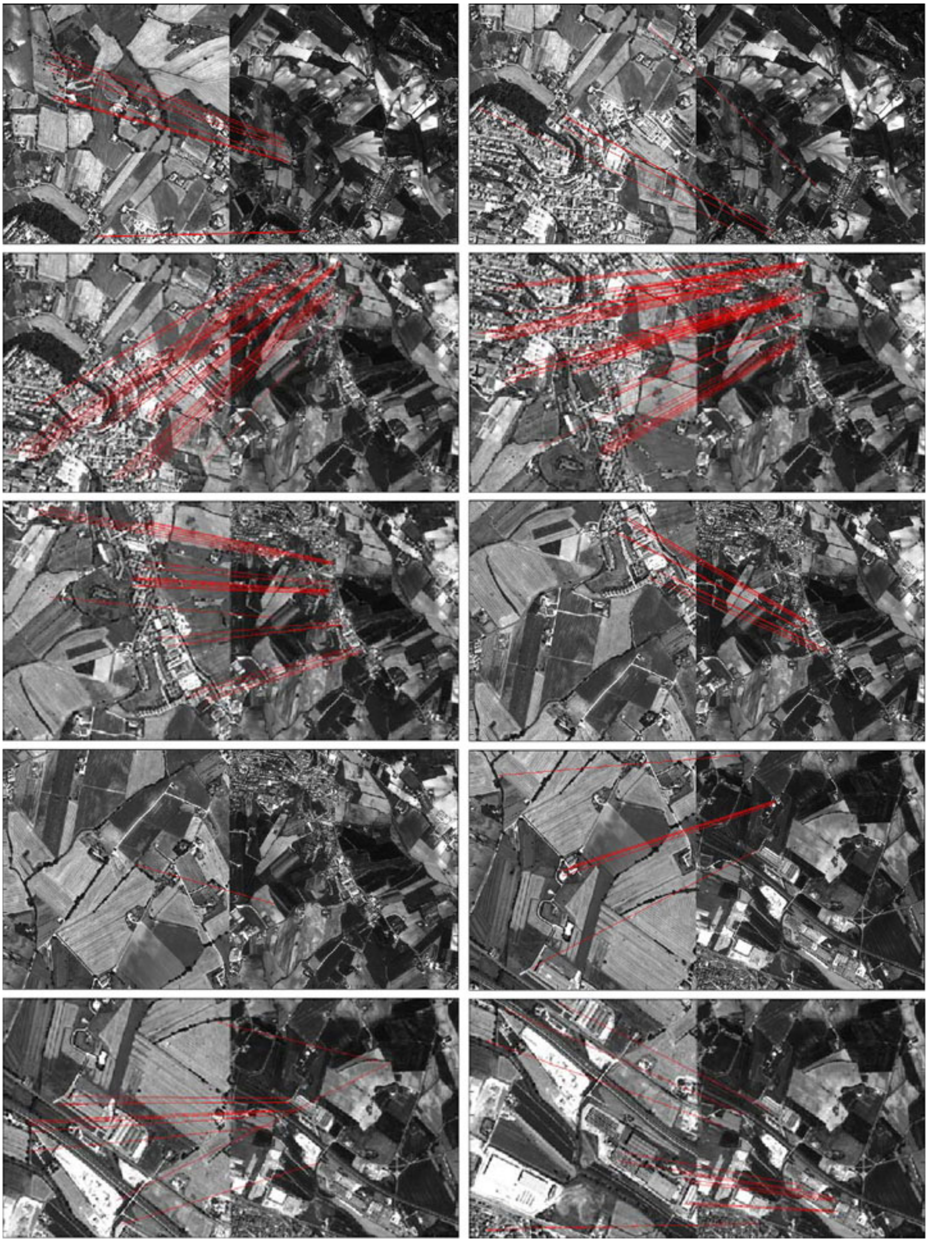

Fig. 8 Some results of the test on absolute localization

Some examples of the matchings computed by the system during the test described above are shown in Fig. 8. Each of the 9 aerial images was compared with 25 images of the map. The system was unable to locate only one image and no false matching was indicated.

Once a correspondence between two images is found, the calculation of the height of flight is based on simple geometric considerations. The distances among the matching points in the first and second image are in the same ratio as the images. This parameter is easily computable. The height of flight and the focal length of the camera related to the reference image are known as well as the focal length of the camera that took the second image. Hence the height of flight of the aircraft can be calculated considering the proportional ratio between the two images.

Although the conclusions resulting from the tests carried out can not yet be considered general, the results obtained by processing the available data show how the proposed system can achieve a quality comparable with that of common low-cost GPS systems in terms of accuracy. The lower frequency achievable in the calculation of the position is compensated by a lower variance of the data.

However the greatest benefit is the smart use of the vision sensor in order to support and overcome the classical limits of the GPS sensor.

The need to keep stored a large set of data and calculation time do not seem to be real limits, considering the rapid development of computing and memory systems nowadays.

5 Conclusions

The vision-based only approach we have implemented to achieve a high accuracy estimate of the position of an autonomous system has been described in this work. During the flight the camera captures images at different positions along the trajectory. A frame-to-frame registration approach is implemented exploiting natural feature points to track images with respect to previous known georeferenced images. An accurate experimental stage on data sets and real world images has permitted to discover, analyze and possibly face different sources of errors that could yield to a position estimation drift.

The high accuracy achieved constitutes a conceptual prove for a vision-based stand alone control system when a high accuracy is required despite using low cost devices.

This method can be also extended to estimate the relative position among several UAVs. Thus, if two UAVs are registering approximately the same scene and it is possible to match a set of features between their images, the relative displacement between the UAVs can be obtained by computing the plane-induced homography matrix that relates their cameras.

References

1. Amidi, O., Kanade, T., Fujita, K.: A visual odometer for autonomous helicopter fight. In: Proceedings of the Fifth International Conference on Intelligent Autonomous Systems (IAS-5) (1998)
2. Ascani, A., Frontoni, E., Mancini, A., Zingaretti, P.: Feature group matching for appearance-based localization. In: IEEE/RSJ 2008 International Conference on Intelligent RObots and Systems—IROS 2008, Nice (2008)

3. Caballero, F., Merino, L., Ferruz, J., Ollero, A.: A visual odometer without 3D reconstruction for aerial vehicles. applications to building inspection. In: Proceedings of the International Conference on Robotics and Automation, pp. 4684–4689. IEEE (2005)
4. Caballero, F., Merino, L., Ferruz, J., Ollero, A.: Improving vision-based planarmotion estimation for unmanned aerial vehicles through online mosaicing. In: Proceedings of the International Conference on Robotics and Automation, pp. 2860–2865. IEEE (2006)
5. Cesetti, A., Mancini, A., Frontoni, E., Zingaretti, P., Longhi, S.: From simulated to real scenarios: a framework for multi-UAVs. In: Carpin, S., Noda, I., Pagello, E., Reggiani, M., von Stryk, O. (eds.) Simulation, modeling, and programming for autonomous robots, Lecture notes in artificial intelligence, vol. 5325, pp. 17–28. Springer, Heidelberg (2008)
6. Cesetti, A., Frontoni, E., Mancini, A., Zingaretti, P., Longhi, S.: A vision-based guidance system for UAV navigation and safe landing using natural landmarks. J. Intell. Robot. Syst. **57**(1–4), 233–257 (2010)
7. Conte, G., Doherty, P.: An integrated UAV navigation system based on aerial image matching. In: 2008 IEEE Aerospace Conference, pp. 1–10, (2008)
8. Corke, P.I., Sikka, P., Roberts, J.M.: Height estimation for an autonomous helicopter. In: Proceedings of ISER, pp. 101–110 (2000)
9. Dickmanns, E.D., Schell, F.R.: Autonomous landing of airplanes using dynamic machine vision. In: Proc. of the IEEE Workshop Applications of Computer Vision, pp. 172–179 (1992)
10. Eisenbeiss, H., Lambers, K., Sauerbier, M.: Photogrammetric recording of the archeoligical site of Pinchango Alto (Palpa, Peru) using a mini helicopter (UAV). Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. **XXXIV-5/C34**, 238–243 (2005)
11. Fischler, M., Bolles, R.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. CACM. **24**(6), 381–395 (1981)
12. Garcia-Pardo, P.J., Sukhatme, G.S., Montgomery, J.F.: Towards vision-based safe landing for an autonomous helicopter. Robot. Auton. Syst. **38**(1), 19–29 (2001)
13. Grant, M.S., Katzberg, S.J., Lawrence, R.W.: GPS remote sensing measurements using aerosonde UAV. In: AIAA 2005-7005, Arlington (2005)
14. Haarbrink R.B., Koers E.: Helicopter UAV for photogrammetry and rapid response. In: 2nd Int. Workshop "The Future of Remote Sensing", ISPRS Inter-Commission Working Group I/V Autonomous Navigation, vol. XXXVI–1/W44 (2006)
15. Haarbrink R.B., Eisenbeiss H.: Accurate DSM production from unmanned helicopter systems. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. **XXXVII**(B1), 1259–1264 (2008)
16. Lowe, D.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. **60**(2), 91–110 (2004)
17. Mancini, A., Cesetti, A., Iaulè, A., Frontoni, E., Zingaretti, P., Longhi, S.: A framework for simulation and testing of UAVs in cooperative scenarios. J. Intell. Robot. Syst. **54**(1–3), 307–329 (2009)
18. Ollero, A., Merino, L.: Control and perception techniques for aerial robotics. Annu. Rev. Control, Elsevier (Francia). **28**, 167–178 (2004)
19. Saripalli, S., Montgomery, J.F., Sukhatme, G.S.: Visually guided landing of an unmanned aerial vehicle. IEEE Trans. Robot. Autom. **19**(3), 371–380 (2003)
20. Saripalli, S., Sukhatme, G.S.: Landing on a mobile target using an autonomous helicopter. In: Proceedings of the International Conference on Field and Service Robotics, FSR (2003)
21. Shakernia, O., Vidal, R., Sharp, C., Ma, Y., Sastry, S.: Multiple view motion estimation and control for landing an aerial vehicle. In: Proceedings of the International Conference on Robotics and Automation, ICRA, vol. 3, pp. 2793–2798. IEEE (2002)
22. Srinivasan, M.V., Zhang, S.W., Garrant, M.A.: Landing strategies in honeybees, and applications to UAVs. In: Springer Tracts in Advanced Robotics, pp. 373–384. Springer-Verlag, Berlin (2003)
23. Vidal, R., Sastry, S., Kim, J., Shakernia, O., Shim, D.: The Berkeley aerial robot project (BEAR). In: Proceeding of the International Conference on Intelligent Robots and Systems, IROS, pp. 1–10. IEEE/RSJ (2002)
24. Wu, A.D., Johnson, E.N., Proctor, A.A.: Vision-aided inertial navigation for flight control. In: Proc. of AIAA Guidance, Navigation, and Control Conference and Exhibit (2005)
25. Yakimenko, O.A., Kaminer, I.I., Lentz, W.J., Ghyzel, P.A.: Unmanned aircraft navigation for shipboard landing using infrared vision. IEEE Trans. Aerosp. Electron. Syst. **38**(4), 1181–1200 (2002)
26. Zhang, Z., Hintz, K.J.: Evolving neural networks for video attitude and hight sensor. In: Proc. of the SPIE International Symposium on Aerospace/Defense Sensing and Control, vol. 2484, pp. 383–393 (1995)